iSSVis: Interactive Visualization of Stroke Subsequences in Table Tennis

Ji Lan, Jiachen Wang, Xinhuan Shu, Kejian Zhao, Jiang Wu, Zheng Zhou, Hui Zhang, and Yingcai Wu

Abstract— Table tennis players usually employ numerous complicated playing techniques and tactics in highly antagonistic, variable, and flexible matches. In-depth analysis of stroke sequences is necessary to obtain insights into the time-varying technical and tactical playing patterns of players. Experienced experts usually watch videos repeatedly to analyze stroke sequences and detect patterns, which is tedious and prone to errors. This paper presents a novel interactive visualization system called iSSVis, which was developed through close collaboration with experts to provide multi-scale comprehensible analysis of diverse stroke sequences of players. The system consists of a matrix-based overview, a glyph-based sequence view, and a figurative detailed view. The intuitive visualization layouts and comprehensible encodings used in iSSVis are highly commended by domain experts, who have identified several new and interesting patterns by using the system. We demonstrate the effectiveness and usability of the system with three case studies and a task-based evaluation. For additional information about iSSVis, readers may visit http://zjuvis.org/ittvis_framework/.

Index Terms—Sports visualization, visual knowledge discovery, sports analytics, visual knowledge representation

1 Introduction

Sports events of table tennis involve high antagonism, tension, celerity, and flexibility. To achieve good results in such circumstances, players utilize a set of complicated techniques and tactics flexibly and adaptively on the basis of various strokes. Analysis of the time-varying features of subsequences picked from all strokes in a match enables analysts to identify temporal characteristics of techniques and tactics of the opponent, as well as significant deficiencies of their player. These findings provide efficient guidance for the future training and competition of athletes.At present, time-oriented analyses of featured strokes are achieved by experienced coaches through repeated observation of videos, which is time-consuming and prone to errors. Statistical analyses and mathematical modeling based on fine-grained data have also been investigated [17, 19, 21, 23, 32, 42, 47]. These methods, which either provide statistics of certain performance indicators or simulate certain structures of a match, follow clear rules and target particular problems. Hence, they cannot provide a holistic view of the variation in a stroke subsequence nor detect unexpected stroke subsequences.

Visualization techniques, which incorporate domain knowledge and human intelligence into the process of data analysis, provide a new possibility for an enhanced analysis of time-varying strokes. Many visualization techniques can identify temporal patterns in one or multiple event sequences [4, 6, 8, 9, 12, 15, 20, 24, 25, 28, 34, 41, 43–45, 48, 49]. However, new difficulties have emerged in visualizing stroke sequences. Examples include providing comprehensible designs for easy access to time-varying attributes of strokes, and presenting data in a manner that matches its practical meaning in table tennis to build a connection between external visualizations and mental models of experts. Existing event sequence visualization techniques cannot be directly applied to these problems, and tailored visualizations are necessary. Our previous system, iTTVis [46], supports the presentation of time-varying patterns of scores, as well as correlation patterns and tactical patterns of stroke attributes. However, analysis of the co-evolution of stroke attributes over a specified stroke subsequence is not well supported. Therefore, the current work is conducted to enable experts to obtain insights into varying features of stroke attributes over diverse stroke subsequences.

Design such a visualization system poses two challenges. First, it poses a challenge to provide a clear overview of diverse stroke subsequences and present the time-varying stroke attributes intuitively over a stroke subsequence. Given the large number and high diversity of existing stroke subsequences, it is difficult to create designs that support presentation and navigation of diverse subsequences. Moreover, each stroke is characterized by three stroke attributes, namely, stroke technique, stroke placement, and stroke position, which possess respective practical meanings. A connection must be established between abstract data and the physical context of table tennis in the designs of these stroke attributes. Therefore, creating intuitive designs for three stroke attributes is another problem. Second, it poses a challenge to visualize the dynamic interplay between the strokes in a subsequence and their nearby strokes, along with the score information and detailed stroke attributes. Each stroke that a player gives is a result of a decision-making process during which the player determines the return details (e.g., stroke technique, stroke placement) on the basis of previous strokes. Experts hope to gain insights into how decision-making processes are implemented, why particular decisions are made, and whether the decisions work. Therefore, the subsequence and the sequences of nearby strokes, along with the decision-making processes between them, must be revealed, which poses a challenge.

We proposed iSSVis which consists of a set of well-organized views to address the challenges. To address the first challenge, we designed a view with informative and well-organized selection matrices. This view provides a clear and holistic presentation of statistical information on diverse stroke subsequences for further selection and analysis. We also designed a glyph-based intuitive timeline view to reveal the varying stroke attributes over a stroke subsequence. This view exploits the metaphors of directions on the table tennis table and thus builds a connection between abstract data and the physical context. To address the second challenge, we introduced a set of juxtaposed flows to allow for a comprehensive and enhanced analysis of the co-evolutionary relationship between key strokes and their nearby strokes. We added winning rates calculated by the Markov chain model to help evaluate the decisions of strokes. The contributions of this study are as follows:

- ◇ Identification of domain requirements for comprehensive analysis of stroke subsequences in table tennis;
- ◇ A visualization system with intuitive glyphs and metaphors that supports navigation of diverse stroke subsequences and comprehensible representation of time-varying features, reasons, and effects of strokes in a subsequence; and
- ◇ Case studies demonstrating the augmented capacity for analyzing time-varying key strokes with the support of our system and new insights obtained for improved understanding of a table tennis match.

- Y. Wu, J. Lan, J. Wang, K. Zhao, and J. Wu are with the State Key Lab of CAD & CG, Zhejiang University. E-mail: {ycwu, lanjizju, wangjiachen, zhaokejian, wu_jiang }@zju.edu.cn. Y. Wu is the corresponding author.
- X. Shu is with the Department of Computer Science and Engineering, Hong Kong University of Science and Technology. Email: xinhuan.shu@connect.ust.hk.
- H. Zhang and Z. Zhou are with the Department of Physical Education, College of Education, Zhejiang University. E-mail:{zhang_hui, zhouzhengzju}@zju.edu.cn.

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

2 Related Work

This section provides an overview of visual techniques in sports visualizations (Section 2.1) and outlines the related event sequence visualizations that inspired our design (Section 2.2).

2.1 Sports Visualization

We classified existing works on sports visualization into three parts according to the layout of visualizations and positioned our approach for visualizing table tennis data.

Location-based visualizations. Layouts of most sports visualizations are based on the physical location. A set of works investigated ball trajectories in a soccer pitch for phase exploration [30], event detection [13], and multi-level abstraction [36]. Another two works involved visual tracking of the trajectories in tennis [3] and baseball [7]. Heatmap was utilized by many sports visualization works [10, 11, 16, 33] to visualize the location-based sports data in real scenarios. Videos have also been included in the analytical process. Parray et al. [27] introduced a form of video visualization to depict and annotate events in snooker video. Director's Cut [37] and Bring it to the Pitch [38] innovatively add visualizations back to the soccer match videos. Designs in these works established a connection between abstract data and the physical context. All the works listed above gave good examples for location-based visual analysis. Table tennis data is inherently location-based as well. In our previous system, iTTVis [46], we used the table tennis table to display location-based characteristics of the data intuitively, and the system was well-received. In this work, we also employed real table tennis tables and used the locations and directions on it to help analysts understand the varying features of stroke subsequences.

Tree-structured visualizations. Several sports applications employ tree visualizations to present tree-structured sports data. Vuillemot et al. [40] and Tan et al. [39] used tree-structured visualizations to present the process of the soccer and basketball tournament. TennisViewer [14] and Game Tree [1] display the tree-structured point progression within a tennis match. These designs illustrate the efficacy of the compact and understandable tree-structured visualizations in the display of tree-structured sports data. A table tennis match involves multi-level data. However, tree visualizations for table tennis data have not been proposed. The current work focused on the time-varying features of strokes instead of the hierarchical information due to the specific requirements of exploring stroke subsequences.

Time-oriented visualizations. Sports data with inherent time-oriented characteristics can also be visualized in a time-oriented layout. A Table [31] and Gap Charts [29] design the rank-based line charts to help understand time-varying rankings of soccer teams. GameFlow [5] and BKViz [22] utilize a series of chronological bars to depict the time-varying statistical metrics of players in a match. These time-oriented visualizations effectively present the time-varying rankings of teams and performance indicators of players. Meanwhile, MatchPad [18] adds a set of intuitive glyphs on the timeline to provide a glimpse of the key events in a real-time rugby match. TenniVis [35] directly arranges intuitive needle gauge-like or ball-like glyphs one by one to compose a time-oriented visualization for a tennis match and game. These designs employ intuitive metaphoric glyphs to represent critical events in sports, thus enabling intuitive analyses. For time-varying table tennis data, the Match View of iTTVis [46] employs a tailored step chart to show the trends and patterns of time-varying scores and relevant information. The Stat View of iTTVis proposes a sequence of matrices to show correlations within sequential strokes. In contrast, the current work employed a time-oriented layout to present the varying features of stroke sequences. Similar to TenniVis [35], the current system proposes a sequence of intuitive glyphs in chronological order. However, our glyphs vary for different domain requirements.

2.2 Event Sequence Visualization

Numerous event sequence visualization works have been proposed. Many of them aimed to detect patterns in multiple event sequences [6, 12, 20, 24, 25, 28, 41, 43–45, 49]. Other works focused on displaying events in one event sequence [4, 8, 9, 15, 34, 48]. LifeLines [34], TimeSlice [48], and Cloudlines [15] directly arrange events horizontally on the timeline in a manner that is straightforward and easy-to-understand. Episogram [4], Eventaction [8], and VariantView [9] pay attention to

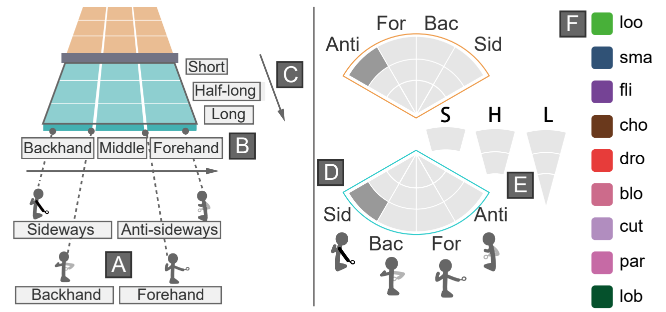

Fig. 1. Three stroke attributes and their encodings: (A) four stroke position values related to different positions where the player hit the ball, (B) and (C) the horizontal and vertical distribution of stroke placement, (D) four stroke position values encoded as four directions, (E) the vertical distribution of stroke placement encoded as the numbers of sectors, and (F) nine stroke technique values encoded by nine colors.

the details of each event in a sequence. Episogram and VariantView use the available vertical screen space to show the social interaction details and gene variant attributes along with the event sequence, respectively. Eventaction displays different types of events in different rows and aggregated events in the same period. However, the case presented in this study, differs from those in previous works. First, our tasks focus on displaying the varying features of a stroke sequence, instead of the temporal distribution. Second, each stroke possesses domain-specific characteristics, such as the stroke and score context. Third, establishing connection between abstract data and the physical context of table tennis is required. We hence designed a tailored glyph-based timeline to display the stroke sequence.

3 Background and System Overview

This section presents the background of the study, data description, requirements, Markov chain model and system overview.

3.1 Background

Table tennis is a sport played by hitting a lightweight ball back and forth across a table by means of a small bat. The action that a player uses the table tennis bat to hit the ball once is defined as a stroke. A table tennis match usually contains hundreds of strokes given by two players. The technical characteristics of each stroke are mainly described by three key stroke attributes, namely, stroke technique, stroke placement, and stroke position. In the sports events of table tennis, the behaviors of players are highly interactive and full of changes. If a player obtains good results using a particular stroke, the opponent would change his (her) corresponding stroke toward this stroke next time. If the player fails to score by using such a stroke several times, he (she) would be less likely to play in this manner. These changing behaviors can be divided into many types according to the variation in attribute values. To better understand different types of changing behaviors, experts find out certain subsequences of all strokes in a match according to certain constraints on the stroke attribute, examining and analyzing the variation in these subsequences. By doing so, experts gain insights into how a player flexibly and adaptively employs various tactics.

3.2 Data Description and Stroke Subsequence

The table tennis data was manually collected from videos of matches, and subsequently organized into a data table. Each row of the table records various attributes of a stroke. The key attributes are as follows:

◇ *Stroke technique:* Nine stroke technique values (Fig. 1F) denote the technical details of the stroke given by a player. Further details can be found on a Wikipedia page [2].
◇ *Stroke position:* Four stroke position values (Fig. 1A) illustrate the position where a player gives the stroke.
◇ *Stroke placement:* Nine stroke placement values defines the drop point of a stroke. They can be grouped into three parts, namely, forehand, middle, and backhand, by horizontal variation (Fig. 1B), or into long, half long, and short by vertical variation (Fig. 1C).

As described in Section 3.1, experts need to find out various stroke subsequences according to certain constrains on the stroke attributes for the analysis of various types of changing behaviors. Formally, **we denote a stroke as a tuple** with nine components describing the three key stroke attributes of the prior, current, and next strokes. Specifically, a tuple $t = (tech_{-1}, pla_{-1}, pos_{-1}, tech_0, pla_0, pos_0, tech_1, pla_1, pos_1)$, where $tech_{-1}, tech_0, tech_1 \in TECH$, $pla_{-1}, pla_0, pla_1 \in PLA$, $pos_{-1}, pos_0, pos_1 \in POS$, and $TECH$, PLA, POS are sets that consist of all attribute values of the stroke technique, stroke placement, and stroke position, respectively. The subscript stands for the number of the stroke (-1, 0, and 1 correspond to the prior, current, and next strokes, respectively). For example, $tech_{-1}$, $tech_0$, and $tech_1$ correspond to the stroke technique value of the prior, current, and next strokes. **We denote a sequence of strokes as a table** (each row is a tuple denoting a stroke). Specifically, a table $R = \{t_1, t_2, t_3 \dots t_n\}$, where t_i is a tuple. The selection of a subsequence from a stroke sequence can be regarded as applying a selection to the table R and producing a new table R' with a subset of R's tuples. The tuples in the resulting table satisfy a certain predicate $F(t)$ (a filter criterion on a tuple t) that involves the nine attributes (Fig. 2). We used the lowercase Greek letter σ to denote selection. The predicate $F(t)$ appears as a subscript of σ.This selection process is written as $\sigma_F(R)$. Specifically, a selection $\sigma_F(R) = \{t | t \in R \wedge F(t) = true\}$. In this manner, various stroke subsequences $R_i = \sigma_{F_i}(R)$ are obtained from the sequence of all strokes in a match, R, according to different predicates $F(t)_i$. We further analyzed them to better understand various changing behaviors of players.

3.3 Requirement Analysis

In our previous work [46], we proposed iTTVis, which supports visual representations of time-varying patterns of scores, as well as correlation patterns and tactical patterns of stroke attributes, for the senior analysts that worked for the Chinese national table tennis team. During the year that experts incorporated iTTVis into their analyses, a set of insights and hypotheses were proposed. However, the experts discovered new problems when browsing the correlation patterns of stroke attributes. The experts could obtain important stroke attribute values from the correlation matrices in iTTVis; strokes with these stroke attribute values were used frequently and with high scoring rates. However, experts hoped to further examine how players vary the use of strokes with these stroke attributes in a match. These details were crutial because they could help experts identify the temporal characteristics of players' techniques and gain insights into the flexible response capability of players. Specifically, experts needed to find out certain stroke subsequences, track their varying features, and gain insights into the reasons and effects of the variation, but these tasks were beyond the capability of traditional methods. Therefore, we were motivated to propose a new method of presenting and exploring the varying features of stroke subsequences of players. We worked closely with experts for eight months. We held weekly meetings to discuss analytical requirements and design principles, developed prototypes, and collected feedback for revision. The major milestones of this process are as follows.

Characterizing problem domains. After using iTTVis for a year, the experts presented the new problems they encountered. To tackle the problems and characterize the initial requirements, we interviewed the experts frequently for three weeks. We aimed to detect varying features of specific stroke subsequences intitially. Then, we extended the requirements to exploring diverse subsequences because the experts also wished to examine unexpected subsequences.

Designing an alpha system. We developed a prototype based on the initial requirements. This prototype displayed the varying features of multiple stroke subsequences without further details on each subsequence. After using our prototype for two weeks, the experts commented that displaying one stroke subsequence at one time would be better and that additional detailed information should be provided for in-depth analyses.

Re-designing a beta system. We collected the feedback and worked on new designs. Through discussions with the the experts for two weeks, we polished the designs and developed a beta system that focuses on flexible navigation and conjoint analysis of an interesting subsequence. Experts were satisfied with this version and incorporated the system into the analysis of varying strokes of players.

Enhancing the beta system. We improved and enriched the beta system iteratively with experts (e.g., displaying stroke subsequences of two players, providing an overview of a stroke subsequence) for

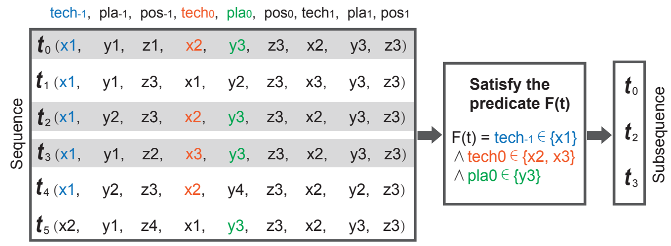

Fig. 2. An example of a subsequence picked out of a sequence.

two weeks. Afterwards, we kept in touch with the experts for further feedback about the system. The summarized requirements as follows:

R1 ***What are the interesting subsequences of all strokes?*** Experts aim to specify proper predicates and select interesting stroke subsequences for analysis. By doing so, experts gain insights into the main features of the temporal tactical behaviors of two players.

R2 ***How do stroke attributes change over a stroke subsequence?*** Three stroke attributes characterize the main technical features of a stroke. Visual tracking of the co-variation of the three stroke attributes in a stroke subsequence is hence of great importance.

R3 ***What are the nearby strokes of the strokes in a subsequence like?*** Nearby strokes of a stroke are defined as the prior two and next two strokes of this stroke and have strong tactical relations with it. Tracking nearby strokes of the strokes in a specified subsequence greatly helps experts understand the variation of this subsequence.

R4 ***What are the effects of selected attributes for each stroke in the subsequence?*** Domain knowledge indicates that each stroke is a result of a decision-making process. A player determines the return details (e.g., stroke technique value, stroke placement value) based on the prior strokes during the process. Experts are curious about how good the selected attribute values and optional values are at each stroke in the subsequence.

R5 ***What are the differences between the selected stroke subsequences of two opposite players?*** Different players have distinct stroke subsequences. Comparing the subsequences of two opposite players helps determine the main differences in their performance and possible reasons for the final results.

R6 ***What are the detailed score and rally context for each stroke in a subsequence?*** The score indicates the overall situation of each stroke. Strokes within the same rally of each stroke explain the use of it and illustrate its effect.

3.4 Markov Chain Model

We employed a modified Markov chain model [32] to calculate changes in winning rates to help evaluate the selected attribute values of each stroke (R4). The model was proposed by Lames [17] and its basic idea is to view each rally as a Markov process representing attribute values randomly changing over strokes. The process begins from an initial state and the final winning rates are obtained when the process is stable.

As shown in Fig. 3, M is an *empirical transition matrix* calculated by all strokes in a match and V is the *initial state probability vector*. The state values include $value_{1,p1}$, $value_{1,p2}$, $\dots$, $value_{n,p1}$, $value_{n,p2}$, $score_{p1}$, and $score_{p2}$, where $value_{1,p1}$, $value_{1,p2}$, $\dots$, $value_{n,p1}$, and $value_{n,p2}$ are all attribute values of a stroke attribute (e.g., stroke technique) of two players. $score_{p1}$ and $score_{p2}$, which represent the states that two players score, are the absorting state values because these two states will not vary further. Each matrix entry presents the transition probability from the row header to the column header. During the computation process, the initial state probability vector is multiplied by the empirical transition matrix repeatedly until probabilities of the non-absorbed state values approach zero (to five decimal points). Then the two winning rates w_1 and w_2 of two players in absorbing vector V_N are attained. These winning rates evaluate players' performance in the entire match. The transition matrix and initial probability vector are derived from the analyzed match. Specifically, $p_{1,2}$ (Fig. 3B) is estimated by the ratio of A to B; A is the number of times that the row header $value_{1,p1}$ transforms into the column header $value_{1,p2}$ and B is the number of times that the row header transforms into all state values

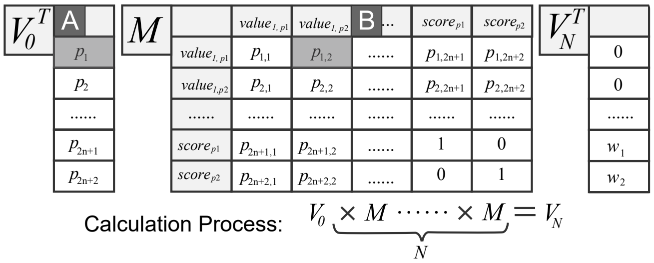

Fig. 3. Illustration of the Markov chain model, where M is the empirical transition matrix and V is the starting vector.

during the match. The p_1 (Fig. 3A) is the frequency the $value_{1,p1}$ appears in the beginning of rallies during the match.

We modified the calculation of transtion matrix M with the help of experts to evaluate the effect of each stroke. The original transition matrix M was calculated with all strokes in a match, and the derived winning rates were used to evaluate the situation of the entire match. We calculated two transition matrices $M_{i\text{-}1}$ and M_i using strokes $t_1, t_2, \ldots, t_{i\text{-}1}$ and strokes $t_1, t_2, \ldots, t_i$, respectively, at each stroke t_i. Moreover, the corresponding winning rates ($w^1_{i\text{-}1}$, $w^2_{i\text{-}1}$) and (w^1_i, w^2_i) were used to evaluate the situations of two players before and after stroke t_i. The amount of changes in the winning rates of the stroker of t_i is $\triangle w^i$ (if the player of stroke t_i is P1, then $\triangle w^i = w^1_i - w^1_{i\text{-}1}$; else, $\triangle w^i = w^2_i - w^2_{i\text{-}1}$). Furthermore, we calculated the transition matrix in view of the three stroke attributes, namely, stroke technique, stroke placement, and stroke position. Accordingly, we obtained three changing winning rates, $\triangle w_{i\_tech}, \triangle w_{i\_pla}, \triangle w_{i\_pos}$.

3.5 System Overview

We developed a well-organized visualization system that supports progressive visual analysis and user-driven exploration. The system consists of three views, namely, a *matrix view*, a *flow view*, and *detailed views*. The matrix view (Fig. 4A) provides an all-around presentation of statistical information on diverse stroke subsequences. The flow view (Fig. 4B) displays varying features of strokes over the selected stroke subsequence and nearby stroke sequences. The detailed views (Figs. 4C, 4D) enables the examination of the context information of any interesting stroke in the sequence. This system contains a set of intuitive components that enhance the understanding of analysts.

The entire system is a web application with three parts. The data preprocessing component extracts the source data from the CSV data tables. The data analysis component employs the model in Section 3.4 to calculate the variation in winning rates at each stroke. In addition, this component calculates the statistical information of stroke subsequences. The visualization component employs D3.js to implement an interactive visualization of the results received from the former part.

4 Visual Design

In this section, we first discuss the design guidelines. Then we illustrate the detailed designs and interactions. Two colors, namely, cyan and orange, stand for two players in the system.

4.1 Design Goals

Existing visualization techniques are unsuitable for comprehensible presentation of the stroke subsequence. We summarized the design goals with the experts as follows:

G1 ***A concise overview and flexible navigation of diverse stroke subsequences.*** There exist numerous and diverse stroke subsequences. Experts ask for effective hints to specify interesting stroke subsequences precisely (R1). The system is hence required to provide an all-around presentation and flexible navigation of the frequency, scoring rates, and varying rates of diverse stroke sequences.

G2 ***Intuitive glyphs to illustrate the stroke attributes.*** Each stroke attribute is deeply connected with its individual practical significance. Experts ask for visual designs with familiar metaphors for intuitive analysis of the variation and features of the three stroke attributes (R2, R3, R5, and R6). Therefore, tailored glyphs and icons for three stroke attributes, which consider respective characteristics, would greatly facilitate expert analysis.

G3 ***Multi-scale display of a selected stroke subsequence.*** Experts wish for the provision of long-term and short-term patterns of a stroke subsequence (R2 and R5). Long-term patterns describe the variation and features of strokes at the match level. They are crucial in describing the overall performance and preference of players. Short-term patterns illustrate the temporary features or changes of strokes at the game or rally level. They are important in detecting interesting individual technical and tactical behaviors.

G4 ***Juxtaposed flows for conjoint analysis of stroke subsequences.*** R3 and R4 request the assistance of nearby strokes and the effects of selected attribute values in analyzing selected strokes. Our work proposes multiple juxtaposed flows that orderly represent the selected subsequence and the nearby stroke sequences, and places relevant effects of the selected attribute values between each two sequences, to allow for a conjoint analysis of stroke subsequences.

G5 ***Comparative analysis of the stroke subsequences between two players.*** A system that enables comparative analysis facilitates understanding of the differences of two players and the final scoring results (R5). A desirable system should provide a comparable overview and details of two players' subsequences.

G6 ***Context unfolding of the interesting stroke on demand.*** Display of a stroke subsequence that allows experts to interact with each stroke directly and obtain the context information immediately is preferred by experts (R6). The system should allow experts to flexibly and quickly examine the required context information of any stroke in the subsequence to identify causes or form hypotheses about the variation and features.

4.2 Matrix View

The system aims to provide a concise and comparable overview and flexible navigation of the statistical information of diverse stroke subsequences to help users locate interesting subsequences (G1, G5).

Justification: We used matrices to present different types of strokes clearly. The matrix view is advantageous in providing a concise overview of multiple subsequences for high information density. Based on well-organized matrices, analysts can quickly browse and compare all types of subsequences and select a desired one. Moreover, experts are familiar with the tabular form, and accept the matrix view well.

Description: The matrix view (Fig. 4A) consisting of seven well-organized matrices with rich interactions fulfills the two tasks: clutter-free presentation and easy navigation of stroke subsequences for two players (G1). The matrices above (Figs. 4A-2, 4A-3, and 4A-4) present statistical information of stroke subsequences specified by pairwise stroke attributes of player one (P1). Similarly, the matrices below (Figs. 4A-11, 4A-12, and 4A-13) present the information of player two (P2). Considering that stroke placement is the most important attribute according to domain knowledge, we placed matrices related to the placement in the middle, and added a matrix (Fig. 4A-10) to present stroke subsequences specified by the stroke placements of two players.

Matrix. Each matrix presents stroke subsequences specified by two stroke attributes. We take the matrix presented in Fig. 5C (amplifying the matrix in Fig. 4A-4) as an example. The row and column where the highlighted matrix entry (Fig. 5C-1) is located correspond to the placement value short middle and position value forehand. The entry thus presents the stroke subsequence of P1 that satisfies the predicate $F(t) = pla_0 \in \{short\ middle\} \wedge pos_0 \in \{forehand\}$. That is, each entry in this matrix indicates a stroke subsequence with the specified values of the placement and position. Three kinds of statistical information are encoded in the entry. The area of the circle encodes the number of strokes in the subsequence. The luminance encodes the scoring rate or varying rate. Matrices in Fig. 4A work in a similar way.

Panel. We provided a selection panel (Fig. 4A-6) to record the attribute values that users specify during the exploration of stroke subsequences. The upside of the selection panel records the attribute values of P1's strokes, whereas the underside records that of P2. For instance, when users click label "ser" (Fig. 4A-1), the label is displayed on the upside of the panel (Fig. 4A-5).

Interactions for exploring diverse subsequences are as follows:

◇ **Switch.** Users can switch the link button to decide whether matrices in the two sides are connected. When this button (Fig. 4A-9) is

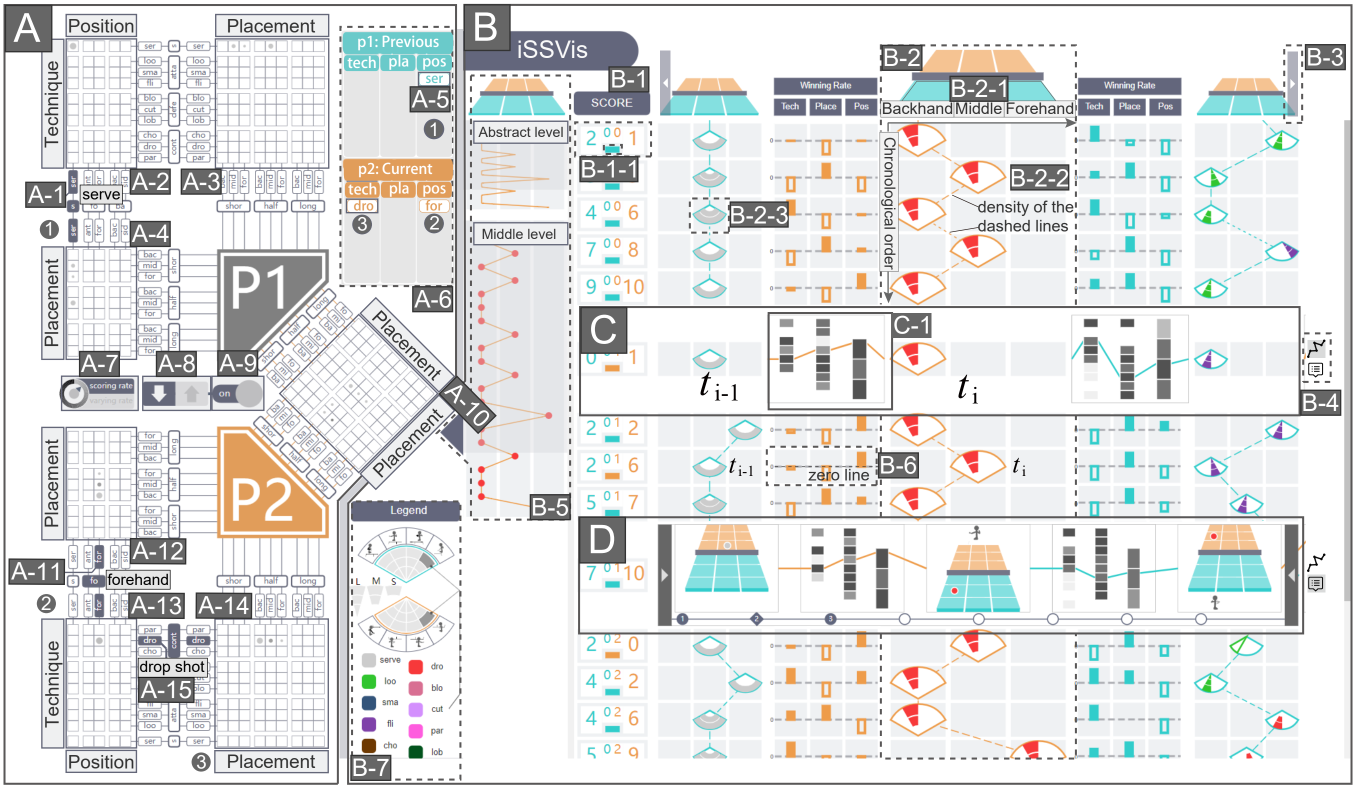

Fig. 4. Interface of iSSVis. The interface comprises three well-coordinated views: the matrix view (including seven well-organized matrices (A-2)-(A-4), (A-10), (A-12)-(A-14), three buttons (A-7)-(A-9), and a panel (A-6)), flow view (including the main stroke flow (B-2) and other flows, an overview (B-5), and a score panel (B-1)), and detailed views (C,D) .

turned off, matrices of each side are separate. When it is turned on, the information of matrices in one side can be filtered by labels in the other side. The arrow button (Fig. 4A-8) further determines the direction. When the arrow points downward, the strokes of P2 are next strokes of P1's strokes. When the arrow points upward, the case is reversed. For instance, the link button is turned on (Fig. 4A-9) and the arrow button points downward (Fig. 4A-8). Therefore, matrices in the two sides are connected, and P2's strokes are the next strokes of P1's. We can also click on button "P1" or "P2" to set P1 or P2's strokes as the strokes for exploration. For instance, clicking on button "P2" in Fig. 4A would select P2's strokes for exploration. We can switch the encoding button (Fig. 4A-7) to decide the encoding of the luminance of circles in matrices.

◇ **Click.** Single-click on the any label are allowed to help users specify stroke subsequences by the attribute value the label stands for. The clicked labels are highlighted and displayed on the panel (Fig. 4A-6), and matrices are filtered accordingly. For instance, we switch the buttons as that the button "P2" is clicked (we select P2's strokes for exploration). Then, matrices in the two sides are connected, and P1's strokes are the prior strokes of P2's. Under this condition, we click on label "ser" in the upside (Fig. 4A-1). Then, the label is highlighted and displayed on the panel (Fig. 4A-5), and all P2's strokes in the underside whose $tech_{-1}$ is not serve are filtered out. At this time, we select P2's strokes that satisfy the predicate $F(t) = tech_{-1} \in \{serve\}$. The matrix view also supports multiple clicks. For instance, we further click label "for" (Fig. 4A-10) in the downside. Then, the label is highlighted and added on the panel. And P2's strokes whose pos_0 is not forehand are filtered out in all matrices. At this time, we select P2's strokes that satisfies the predicate $F(t) = tech_{-1} \in \{serve\} \wedge pos_0 \in \{forehand\}$. Similarly, after we click on label "dro" (Fig. 4A-14), this label is displayed on the panel and now we select P2's strokes that satisfies the predicate $F(t) = tech_{-1} \in \{serve\} \wedge pos_0 \in \{forehand\} \wedge tech_0 \in \{drop\ shot\}$. Specifically, we click on the labels according to the statistical information presented on the matrix, and after each click this information are filtered and updated. The subsequence consisting of selected strokes will be displayed in the flow view.

◇ **Group and Hover.** Users can group multiple attribute values by ctrl + clicking several labels. Therefore, we support the specification of a subsequence by a group of attribute values. We can hover on any matrix entry or label, and the corresponding rows and columns will be highlighted. In addition, when users hover on any label or matrix entry in the side of one player, the same label and matrix entry in the side of the other player will be highlighted (G5).

4.3 Flow View

The system needs to provide tailored glyphs and juxtaposed flows for familiar and conjoint analysis of the stroke subsequence specified in the matrix view (G2, G4). Besides, multi-scale display and comparative analysis are also required (G3, G5).

Justification: We used a set of efficient visual channels to encode the varying stroke attributes over the stroke subsequences. Given that the variation in stroke placement in the horizontal direction and the variation in stroke techniques are the most significant time-varying features [26], we used outstanding channels, namely, the position (Fig. 4B-2-1) and color hue (Fig. 1F), to represent them in the flow view. Meanwhile, the length and angle, which are also highly ranked channels [26], are employed to represent the variation in stroke placements in the vertical direction (Fig. 1F) and the variation in stroke positions (Fig. 1D). These encodings are good metaphors for the practical meaning of the encoded data and all of them are well coordinated as a glyph. Furthermore, we juxtapose two prior strokes, the current stroke, and two next strokes for each stroke in the specified stroke subsequence (Fig. 5A). In this manner, we can determine the reasons for and the effects of the varying features over the stroke subsequence from the most relevant strokes.

Description: The flow view consists of five stroke flows (Fig. 5A) and four bar chart flows (Fig. 5B) with a set of tailored glyphs. The stroke flows present stroke sequences and the bar chart flows present effects of decisions between two stroke sequences. The main stroke flow (Fig. 5A-3) is placed in the middle; it presents the selected stroke subsequence. According to domain knowledge, three consecutive strokes make up a tactic. Each selected stroke can form a tactic with two

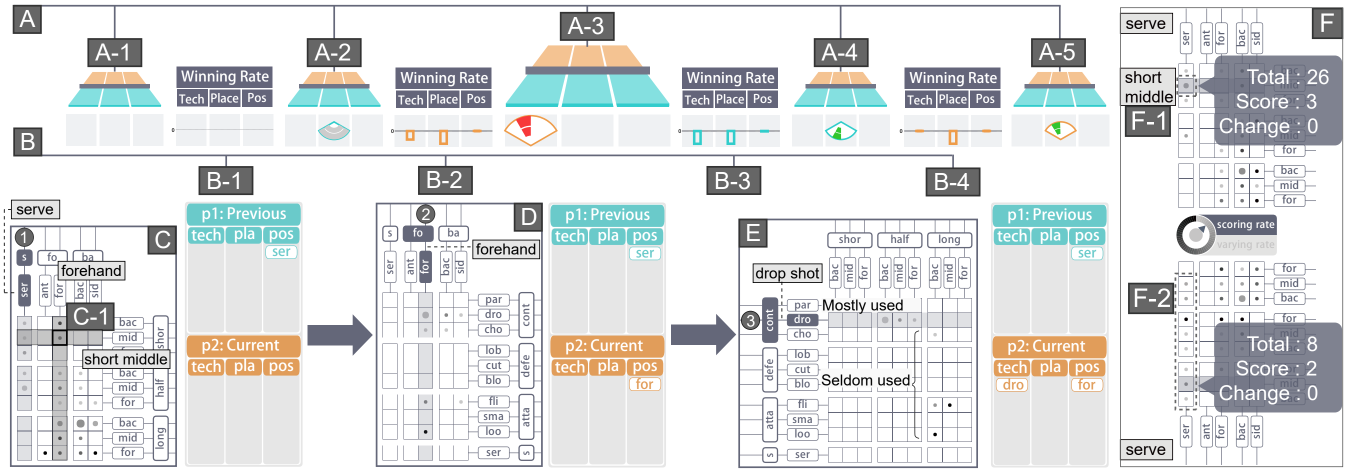

Fig. 5. (A) and (B) show five stroke flows and four bar chart flows in the flow view, respectively. (C),(D), and (E) show the analysis pipeline of the first case study. They amplify matrices presented in Fig. 4A-4, Fig. 4A-12, and Fig. 4A-13, respectively. (F) shows matrices presenting two players' stroke subsequences specified by the stroke placement and stroke position.

prior strokes, one prior stroke and one next stroke, or two next strokes. Therefore, two small stroke flows representing the two prior (Figs. 5A-1, 5A-2) or two next strokes (Figs. 5A-4, 5A-5) are placed on each side of the main flow (G4) to show the tactical patterns. Bar chart flows (Fig. 5B) are placed between two adjacent stroke flows (G4). These charts are used to evaluate the effects of the decisions on stroke attributes of the next strokes in the cases of the prior strokes. We provided five stroke flows and bar chart flows under three modes, with certain consecutive three of the five flows displayed in each mode. For instance, we display the prior, current, and next stroke flows in Fig. 4B. The switch button (Fig. 4B-3) allows switching between different modes. With these designs, experts can quickly detect the main varying features of the selected strokes and identify possible reasons from the variation in their nearby strokes.

Stroke flow: We employ a set of visual designs to encode the three stroke attributes in the stroke flow (G2). These intuitive designs connect abstract data to its practical meaning, thus allowing a comprehensible analysis of the variation in the three stroke attributes.

◇ **Horizontal variation in stroke placement.** Fig. 1B shows the horizontal and vertical distributions of stroke placement values. According to the experts, the variation in horizontal direction is more significant than that in the vertical direction. Therefore, we used the horizontal side of the table as the x-axis of the stroke flow to vividly and outstandingly present the horizontal variation (Fig. 4B-2-1). The other direction of the stroke flow, the y-axis, was used to denote the chronological order.

◇ **Stroke position.** Fig. 1A shows four distinct stroke position values related to four positions where the player hit the ball. The ball thus comes from different directions given the different positions. Therefore, we employed two opposite fans to represent strokes of two players and employ four directions within each fan to represent four stroke position values (Fig. 1D). Besides, the experts commented that the direction in each glyph shows the varying features over the sequence very well. As serve strokes do not have the stroke position, we use the glyphs without direction information to represent them (Fig. 4B-2-3). The density of dashed lines (Fig. 4B-2-2) connecting two fans encodes the time span between the two fans.

◇ **Vertical variation in stroke placement and stroke technique.** As shown in (Fig. 1C), the stroke placement values are divided into three groups in the vertical direction indicating the distance from the middle of the table to the drop point of the ball. We hence used the number of the sectors to encode this vertical distance (Fig. 1E). The color of the sector encodes the stroke technique (Fig. 1F).

An overview (Fig. 4B-5) that simplifies the main stroke flow is provided (G3). At the abstract level, the x-axis is used to encode the horizontal distribution of stroke placement. At the middle level, balls with different colors encodes strokes with different technique values. The overview is coordinated with the flows. The score panel (Fig. 4B-1) shows the corresponding score information. The color of the rectangle under the scores represents the winning player of the rally.

Bar chart flow. The bar chart flow is placed between two stroke flows. A bar chart (Fig. 4B-6) connects the prior stroke t_{i-1} and the current stroke t_i. It presents the changes in three winning rates, namely, $\triangle w_{i\_tech}, \triangle w_{i\_pla}, and \triangle w_{i\_pos}$ of the stroker after stroke t_i (as illustrated in Section 3.4). These changes are encoded as the directions (positive upward and negative downward) and lengths (absolute value) of three bars. A zero line indicating no changes is placed in the middle. These bar charts help detect good or bad decisions about stroke attributes of the current strokes in the cases of the prior strokes.

Interactions on the flow view are as follows:

◇ **Switch one/two players.** When the link button (Fig. 4A-9) is enabled, the flow view only displays strokes of one player (Fig. 6). When the link button is disabled, the flow view displays strokes of two players (Fig. 8) (G5).

◇ **Switch the mode.** We can switch among three modes by clicking the switch button (Fig. 4B-3) to examine the first three, middle three, and last three of the five flows (Fig. 5A).

◇ **Switch the viewpoint.** We can click on the table icon to switch the viewpoint. The flabellate glyphs then reverse their directions.

4.4 Detailed Views

The system needs to support context unfolding of any stroke in the stroke sequence (G6). That is, the rally context and decision-making processes should be explored further on demand.

Description: We provided decision-making views (Fig. 4C) and rally views (Fig. 4D) to help analysts understand the decision-making process and rally context of the specified stroke, respectively (G6). The two views are displayed when analysts click on the icons (Fig. 4B-4).

Decision-making view. As shown in (Fig. 4C-1), the decision-making view connects the prior stroke t_{i-1} and the current stroke t_i. This view is the detailed view of the bar chart (Fig. 4B-6) and reveals the decision process of the hit player at the current stroke t_i. As discussed in Section 5, we calculate the amounts of changes in the three winning rates $\triangle w_{i\_tech}, \triangle w_{i\_pla}, \triangle w_{i\_pos}$ of the stroker after the current stroke t_i. These amounts of changes are caused by selecting the three attribute values of the stroke t_i. We used three rectangles orderly linked by a line to represent selected attribute values of stroke technique, placement, and position of stroke t_i, respectively (Fig. 4C-1). Their luminance represents the amounts of changes in the winning rates caused by the three selected attribute values. In contrast, the luminance of the other rectangles beside the three linked rectangles encodes the amounts of changes in the winning rates when the player selects other attribute values (also calculated by the Markov chain). A few attribute values are impossible to use in the current stroke based on domain knowledge, and their rectangles are hence hidden.

Rally view. The rally view (Fig. 4D) provides all strokes in the rally which the specified stroke belongs to. The position and color of the ball on the table represent the stroke placement and technique, and

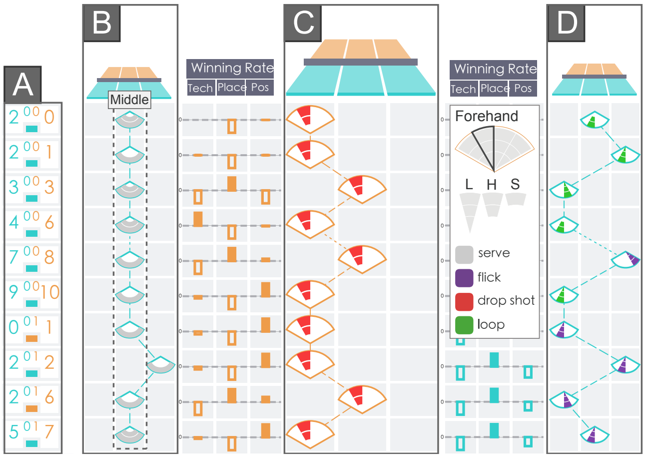

Fig. 6. Analysis pipeline of the case study discussed in Section 5.1.1 (A) shows the accompanying score contexts of the strokes. The colors of the small rectangles under the scores indicate the winners of the rallies. (B) shows the sequences of the prior strokes. (C) shows the stroke flow of the subsequence further filtered by the technique value drop shot. (D) shows the sequences of the next strokes.

the figure beside the table encodes the stroke position. The decision-making processes are also placed between every two strokes. This view displays three strokes at one time, thereby showing a complete tactic. A timeline is added to provide a navigation into strokes of this rally.

5 Case Studies and Discussion

This section presents three case studies and a user study. These studies were conducted on the Google Chrome on a PC (equipped with Intel Xeon E3, 32GB of memory, and a 1920*1080 display). We invited two domain experts A and B to use the system for a week. The experts were sports analysts with a focus on techniques and tactics of table tennis players. Specifically, the expert A worked for the Chinese national table tennis team. The expert B was a PhD student majoring sports science. He was also a table tennis athlete and senior analyst on table tennis data. We gave a tutorial on how to use the system for the experts, answered their questions about the system when they used it, recorded their exploration process, and collected their feedback afterward.

5.1 Case Studies

We worked with expert B to discuss the insights and determine how the system was used to discover patterns. The results are summarized as three case studies.

5.1.1 Low-equality Drop Shot Causes Losing

Wang Hao (cyan, P1) and Ryu Seung-min (orange, P2) on July 26, 2007. The expert was interested in the characteristics of Ryu's strokes next to the serve strokes. He firstly turned on the link button (Fig. 4A-9) (connecting strokes of two players), clicked the button "P2" (choosing Ryu's strokes for exploration), and turned the arrow downward (Fig. 4A-8) (setting Wang's strokes as the prior strokes to Ryu's). He then clicked label "ser" in the side of Wang (Fig. 5C) (amplifying the matrix (Fig. 4A-4)). In this way, he selected the strokes of Ryu next to the serve strokes of Wang. He then browsed the matrix (Fig. 5D) (amplifying the matrix (Fig. 4A-12)) that presents Ryu's strokes specified by position and technique. He quickly found that Ryu used forehand mostly because of the many large points in the column of label "for". He clicked label "for" and browsed the matrix (Fig. 5E) (amplifying the matrix (Fig. 4A-13)). He found that Ryu mostly returned the serve ball with the technique drop shot. Other techniques, such as flick, loop, and chop long, were seldom employed. The expert commented that from his previous experience, a player was likely to use flick, drop shot and chop long flexibly to receive a server stroke. However, Ryu behaved differently against prior knowledge because he was prone to employing control techniques, such as drop shot, instead of the offensive technique flick. The reason that Ryu was not good at flick might account for it. The expert thought that this view showed the typical characteristics of Ryu' strokes next to the serve strokes. The expert wanted to know more varying features of the strokes with drop shot. And he clicked label "dro" (Fig. 5E). At this time, the experts selected Ryu's strokes that satisfy the predicate $F(t) = tech_{-1} \in \{serve\} \wedge pos_0 \in \{forehand\} \wedge tech_0 \in \{drop\ shot\}$. Then the flow view displayed all Ryu's strokes next to serve strokes with forehand and drop shot (Fig. 6).

The expert explained that the forehand strokes showed dominant advantages of a player. He examined the strokes prior to these strokes to find the reasons for employing forehand. He found that Wang struck the ball to the middle hand area at most serve strokes (Fig. 6B), thus providing good chances for Ryu to employ forehand. However, most strokes that Ryu gave were to half-long area (Fig. 6C), which indicated that Ryu did not control the ball within the short area successfully. As the stroke flow next to these strokes shows, Wang mostly used offensive techniques, such as loop (green) and flick (purple), to respond (Fig. 6D), thereby taking advantage of the good offensive opportunities after the poor control strokes of Ryu. The expert also noticed that the placements of these offensive strokes by Wang varied frequently and irregularly in the horizontal direction (the horizontal positions of strokes vary frequently). In contrast, the techniques employed did not vary so frequently (the color of strokes is relatively stable). The score panel (Fig. 6A) showed that Wang won most rallies after these strokes (the color of most rectangles under the scores is cyan). The expert said that compared to the placement, Wang did not vary the technique so frequently when receiving Ryu's drop shot strokes and obtained effective results. These offensive strokes deserve further study.

This study demonstrated that the expert could flexibly locate the required stroke subsequence through the well-organized the matrix view, quickly detect the main varying features, and deduce relevant reasons and results through the multi-scale and juxtaposed flows.

5.1.2 Adjusting Flick with Anti-sideways Improves performance

Wang Hao (cyan) and Patrick Baum (orange) on May 2, 2012. The expert wanted to explore the strokes of Wang in terms of stroke technique and position. He examined the matrix that presents stroke subsequences specified by the stroke technique and position for Wang, and found that the scoring rates (Fig. 7A-1) and varying rates (Fig. 7A-2) (denoted by the luminance of the circle) presented in the rows of loo and fli were relatively high. As the expert was familiar with the strokes using loop by Wang, he clicked on label "flick" and explored how the strokes using flick varied. He found from the flow view that Wang mostly employed flick with anti-sideways (Fig. 7B). The expert further selected strokes with anti-sideways by clicking label "ant".

The expert found a distinct trend of the placements from the backhand area to the forehand area (Fig. 7E-1). Combining the score information, he deduced that Wang struck the ball to the backhand area in the first game but obtained bad results. In the second game, Wang struck the ball more to the forehand and middlehand area and won more rallies. After losing a rally when striking the ball to the middlehand area (Fig. 7F), Wang kept striking to the forehand area and obtained positive results. The expert also browsed all the bar charts prior to the main stroke flow (Fig. 7D) and found that the bars in the pos (position) columns are all above the zero lines. These bars present the differences in the winning rates of Wang caused by the position of strokes in the main flow. That the bars are all above the zero lines illustrated that the decisions of Wang to employ anti-sideways had good effects on the winning rates. The expert commented that Wang featured his effective use of flick and anti-sideways.

The expert picked a stroke (Fig. 7C) with the score four to three in the third game to examine how Wang employed anti-sideways to win such a key rally. As commented by the expert, the rally with the score four to three is important because the game enters the middle phase after it. During this rally (Fig. 7G), Wang returned the serve stroke with flick and anti-sideways at the second stroke (Fig. 7G-2). After the opponent returned the stroke using loop to long middle hand at the third stroke (Fig. 7G-4), Wang quickly moved to the backhand area and struck the ball with sideways at the fourth stroke (Fig. 7G-6). The opponent, Baum, who did not expect that Wang would use sideways, failed to return the ball at the last stroke and lost (Fig. 7G-7). The expert further examined the decision-making processes. In the decision-making views prior to the second, third, fourth stroke (Fig. 7G-

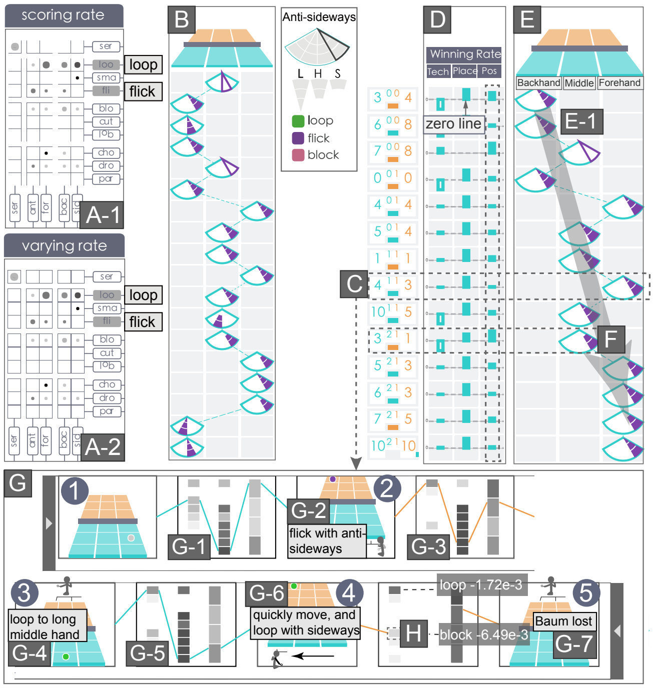

Fig. 7. Analysis pipeline of the case study discussed in Section 5.1.2. (A-1),(A-2) show the matrix presenting the scoring rate and varying rate of stroke subsequences specified by the stroke technique and position of Wang. (B) shows the main stroke flow of a subsequence of strokes by Wang specified by the technique value flick. (C) shows a stroke picked for further analysis in the rally view. (D) shows the bar chart flow prior to the main stroke flow. (E) shows the main stroke flow of a subsequence of strokes by Wang specified by the technique value flick and a position value anti-sideways. (F) shows the last stroke in the match in which Wang struck to the middlehand area with flick and anti-sideways. (G) shows each stroke in the rally of the selected stroke. (H) shows the amount of changes caused by stroke technique loop and block.

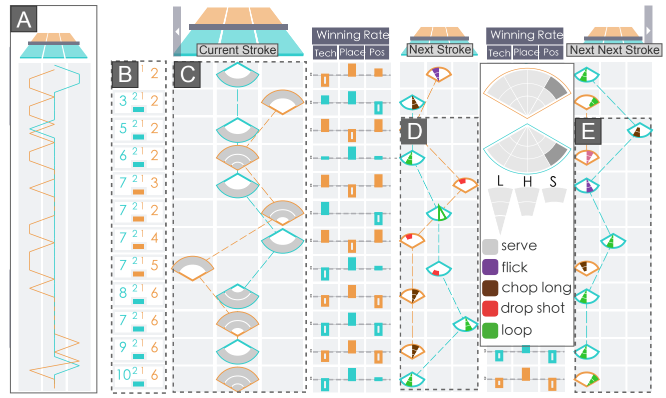

Fig. 8. Analysis pipeline of the case study discussed in Section 5.1.3. (A) shows the abstract level overview of the serve stroke subsequences of two players. (B) shows the accompanying scores. (C) shows serve strokes of two players. (D) shows the sequences of the strokes next to the serve strokes of two players. (E) shows the sequences of the strokes next next to the serve strokes of two players.

1, G-3, G-5), the luminance of linked rectangles is relatively dark, which indicated relatively good effects of these three strokes. In the decision-making view prior to the last stroke, however, the luminance of a linked rectangle representing the selected technique value is relatively lighter (Fig. 7H). This indicated that a better attribute value existed in the last stroke. The expert hovered on the rectangle and the darkest rectangle and found that they are block and loop, respectively. Loop might be a better choice than block at the last stroke.

This study demonstrated that the matrix view could provide effective hints for the expert to specify interesting stroke sequences, and the detailed views could immediately unfold the context of any stroke in the subsequence displayed on the flow view.

5.1.3 Controlling Serve Length Facilitates Attack

Wang Hao (cyan) and Ryu Seung-min (orange) on February 17, 2009. The expert mentioned that the serve stroke was the beginning of the serve tactic and had an important impact on the entire rally. He thus compared the serve stroke of two players in this match. He firstly turned off the link button (Fig. 4A-9) (strokes of two players are separate) and browsed the two matrices representing stroke subsequences specified by the stroke placement and stroke position (Fig. 5F). The expert compared the distributions of serve strokes over different stroke placement values of two players. He found that most of the serve strokes of Wang were struck to short middlehand (Fig. 5F-1). In contrast, the serve strokes of Ryu were evenly distributed over different stroke placements (Fig. 5F-2). He further hovered on the matrix entry specified by serve and short middlehand for Wang, and the tag showed that this subsequence contained 26 strokes. The tag of the corresponding entry of Ryu showed that the subsequence of Ryu only contained eight strokes.

Then the expert browsed the flow view to explore the variation in all serve strokes of Wang and Ryu. From the abstract-level overview, the expert found that the stroke placements of serve strokes for Wang were always the middlehand in the middle part of the match (Fig. 8A) (the cyan line is basically straight in the middle column). The serve strokes of Ryu, however, varied frequently between backhand and middle hand (Fig. 8A). As the expert mentioned, the stroke placements of the serve strokes of Wang and Ryu were stable and variable, respectively. The expert further explored the detailed time-varying features in the stroke flows. He quickly found from the main stroke flow that most of the serve strokes of Wang were short strokes and those of Ryu were usually half-long strokes (Fig. 8C). The expert commented that Wang was better in controlling the length of serve strokes than Ryu. The expert switched the tactic modes to examine the next two strokes of the serve strokes. The expert detected that Ryu had to employ control techniques, such as chop long and drop shot, at the second stroke due to Wang's short serve strokes (Fig. 8D). Then Wang grasped the chance to attack first with loop in the third stroke (Fig. 8E). In contrast, Wang directly attacked with loop at the second stroke because Ryu did not control the ball within the short area (Fig. 8D). According to the score information on the left, Wang won more rallies (Fig. 8B).

This study demonstrated that the matrix view and flow view could support comparison of subsequences of two players.

5.2 Domain Expert Feedback

Overall, the expert thought our system provided great assistance in exploring and analyzing various interesting stroke sequences. Specifically, the experts appreciated two aspects of our system. First, the experts approved of the stroke flow because of its familiar metaphors and figurative icons. They particularly appreciated the way that the flow view aligned the horizontal direction of the table icon with the x-axis of the flow. Expert A commented, "The way the flow view presents strokes matches well with the way we players look at the ball. It feels like the strokes flow out of the table. This view commendably brings reality into data analysis." Second, the experts appreciated the presentation of the winning rates along with the practical features of strokes. The Markov model has been proposed to measure objective winning rates employed table tennis [17, 32]. The system further enables experts to combine objective winning rates with the subjective evaluation of the stroke attributes. Expert B commented, "I think the Markov chain model and the visual display of the puzzling winning rates along with the familiar stroke attributes are meaningful". Two suggestions were proposed by experts. First, the system should be improved to enable analysis of table tennis doubles. Second, the decision-making view, which presents the decision-making process of a stroke, can be further extended to present the decision-making chain of a rally.

5.3 Task-based User Study

We conducted a task-based user study to evaluate the effectiveness and usability of our system. We invited 16 volunteers (eight females and eight males) to complete a series of tasks. The participants were undergraduate and graduate students from different majors and had experience in playing table tennis and data analysis. The match between Hao Wang and Ryu Seung-min on July 26, 2007, was used for the study.

Tasks and Design. We developed seven tasks (Table 1) which are multiple-choice questions based on the requirements (Section 3.3). To ensure the objective assessment of the participants regarding the usability of the system, we also designed a questionnaire (Table 2) with a seven-point Likert scale (1 means strongly disagree, and 7 means strongly agree). The first six questions were based on the design goals (Section 4.1), and the last three questions were for general evaluation.

Procedure. The study began with a brief introduction of table tennis terminologies and our table tennis data. A tutorial was provided to help participants understand the visual designs in the system. We explained the designs and functions of the system with examples and showed the participants how to use the system. Each participant was then asked to accomplish seven tasks (Table 1) by using iSSVis. Instructors would be aside to ensure accurate comprehension of the tasks and designs of the system. We calculated the pass rate of each task on the basis of their answers. Afterward, the participants were asked to evaluate our system through the post-study questionnaire (Table 2). Finally, we interviewed each participant to elicit further feedback. The entire procedure lasted for around 40 minutes for each participant.

Table 1. Experiment tasks

T1	What is the stroke subsequence with the most variation in the matrix specified by the stroke placement and position of Player 1?	R1
T2	What is the horizontal distribution of the stroke placement over the subsequence consisting consisting of strokes using loop by Wang?	R2
T3	What are the varying features of the subsequence consisting of strokes using block by Ryu?	R2
T4	What is the variation in the techniques over the strokes next to the strokes in the subsequence mentioned in the last task?	R3
T5	Examine the fourth stroke of Wang using loop. Is the chosen stroke technique value the best choice according to the winning rates ?	R4
T6	What are the varying features of the two subsequences consisting of strokes using drop shot by two players?	R5
T7	Examine the second stroke of Wang's strokes using drop shot. What is the detailed rally information of this stroke like?	R6

Table 2. Post-study questionnaire.

F1	The matrix view provides a concise overview of diverse stroke subsequences.	G1
F2	The flow view contains intuitive glyphs to illustrate the stroke attributes.	G2
F3	The flow view supports a multi-scale display of the stroke subsequence.	G3
F4	The juxtaposed flows help in conjoint analysis of stroke subsequences.	G4
F5	The flow view supports comparative analysis of two players' strokes.	G5
F6	The detailed views provide useful context information.	G6
F7	The interface is aesthetic and has a consistent color encoding.	
F8	The system consists of well-coordinated views and user-friendly interactions.	
F9	The system can help users perform an in-depth analysis of diverse stroke subsequences.	

Results. Fig. 9 shows the pass rates of the tasks and the mean values and 95% confidence intervals of the post-study questionnaire ratings.

- ◇ **Pass rates of the tasks.** The result (Fig. 9A) shows that except for tasks 2 and 7 with the pass rates of 81%, other ratings are higher than 93%. Overall, the result with the average pass rate of 92.9% demonstrates the effectiveness of our system. The pass rate for T2 is lower because a few common users confuse the backhand area with forehand area. The passing rate for T7 is lower because a few common users could not understand the details of a specific stroke.
- ◇ **Post-study questionnaire ratings.** Overall, the average rate is 6.18. Except for those of F2, F6, and F7, all rates are higher than 6 (Fig. 9B). This result demonstrates that users generally agreed on the usability and effectiveness of the system. Furthermore, F2 was rated low. We found from the user feedback that a few users were unfamiliar with the stroke attributes of table tennis and thus did not accept the encodings very well. The rates of the last three questions, namely, F7, F8, and F9, are all higher than 5.9, which proves that users approved of the aesthetics and effectiveness of our system.

User Feedback. In the post-study interview, most participants agreed that the system was well designed with an intuitive and easy-to-understand flow view, an effective and informative matrix view, and a set of comprehensible detailed views. A user stated, "When I first saw the system, I thought it was informative and cool. With the help of your introduction, I quickly mastered the system to find various information." Several users also proposed that the matrix view could be further simplified despite its smooth interactions.

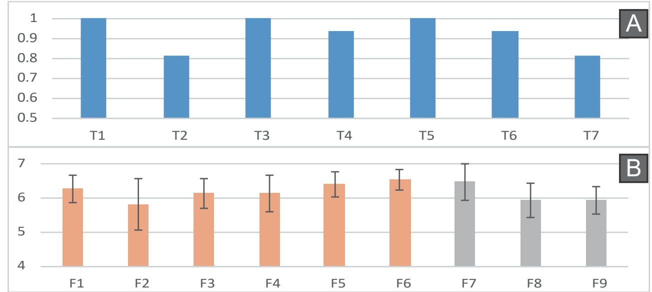

Fig. 9. (A) Pass rates of the experiment tasks. (B) User ratings on each part of the systems in the post-study questionnaire.

5.4 Discussion

Lessons Learned. We learned two main lessons from the entire design process. First, the stroke subsequence in table tennis should be interpreted and analyzed in the context of nearby strokes. As table tennis matches are highly interactive and antagonistic, each stroke is highly related to nearby strokes. We hence designed the juxtaposed stroke flows for conjoint analysis. We learned from this work that a sequence could be visually analyzed in the context of relevant sequences. This condition inspired us to think about an event sequence data from a new perspective. We could define certain predicates to select event subsequences from an event sequence and analyze these event subsequences along with the sequences of nearby events. Second, visual encodings of sports data should be closely connected with the physical context. Data visualization with familiar metaphors provides a possibility for sports experts to fuse interesting data patterns with their practical significance. In this work, the horizontal distribution of the stroke placement corresponds to the horizontal direction of the table tennis table. We hence employed the horizontal direction of the table icon as the x-axis of the stroke flow to enhance the understanding of experts.

Limitations. iSSVis has two limitations. First, the system cannot support comparisons of multiple stroke sequences of many players within a set of matches. It also does not support the analysis and tracking of the stroke sequences of a player in several matches over a long time. We plan to enrich the system with this function in the future. Second, the system does not support explorations of the winning rates calculated by the Markov chain model. Although the model has already been employed in table tennis and simulates a match well, our experts could not understand the computational procedure and gain insights into intermediate results. We plan to introduce new visualizations to help interpret the model.

General Applicability. iSSVis can be extended to support the analysis of stroke sequences in other single-player sports with similar game mechanisms and data structures, such as tennis and badminton.

6 Conclusion

This work investigated the problem of analyzing diverse stroke subsequences. We worked with experts closely to identify major problems in analyzing featured stroke subsequences and developed a visualization system to enhance the analysis. We conducted three case studies and one user study to evaluate our system. The main implications of this work are as follows. First, this study presented the problem of visualizing stroke subsequences in table tennis. We identified the requirements for intuitive and conjoint analysis of the respective strokes attributes, and provided a set of heuristic visual designs, such as intuitive glyphs and juxtaposed flows. We hope our work could encourage other visualization practitioners to study this problem. Second, our system allows experts to determine the main varying technical features of table tennis players in a match and detect insightful patterns for further research. Compared to current methods, which are limited to video analysis, our approach enables quick location of featured stroke subsequences, effective exploration of a stroke subsequence, and interactive pattern unfolding of the variation. With this technical assistance, the experts have completed effective and insightful analysis of several table tennis matches, and detected interesting patterns. We plan to further explore visual comparisons of multiple stroke subsequences and visual tracking of the variation in stroke subsequences of a player in a set of matches.

References

[1] Game tree: Point progression. http://tennisviz.blogspot.jp/2015/10/game-tree.html.

[2] Types of strokes. https://en.wikipedia.org/wiki/Table_tennis#Types_of_strokes.

[3] A 3D lesson in clutch point serving by S.Stakhovsky — GameSetMap. http://gamesetmap.com/?p=182, 2013.

[4] N. Cao, Y. Lin, F. Du, and D. Wang. Episogram: Visual summarization of egocentric social interactions. *IEEE Computer Graphics and Applications*, 36(5):72–81, 2016.

[5] W. Chen, T. Lao, J. Xia, X. Huang, B. Zhu, W. Hu, and H. Guan. GameFlow: Narrative visualization of NBA basketball games. *IEEE Transactions on Multimedia*, 18(11):2247–2256, 2016.

[6] Y. Chen, P. Xu, and L. Ren. Sequence Synopsis: Optimize visual summary of temporal event data. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):45–55, 2018.

[7] C. Dietrich, D. Koop, H. T. Vo, and C. T. Silva. Baseball4D: A tool for baseball game reconstruction &visualization. In *Proceedings of IEEE Conference on Visual Analytics Science and Technology*, pages 23–32, 2014.

[8] F. Du, C. Plaisant, N. Spring, and B. Shneiderman. EventAction: Visual analytics for temporal event sequence recommendation. In *Proceedings of IEEE Conference on Visual Analytics Science and Technology*, pages 61–70, 2016.

[9] J. A. Ferstay, C. B. Nielsen, and T. Munzner. Variant View: Visualizing sequence variants in their gene context. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2546–2555, 2013.

[10] A. Franks, A. Miller, L. Bornn, and K. Goldsberry. Counterpoints: Advanced defensive metrics for NBA basketball. In *Proceedings of MIT Sloan Sports Analytics Conference*, 2015.

[11] K. Goldsberry. CourtVision: New visual and spatial analytics for the NBA. *Proceedings of MIT Sloan Sports Analytics Conference*, 2012.

[12] D. Gotz and H. Stavropoulos. DecisionFlow: Visual analytics for high-dimensional temporal event sequence data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1783–1792, 2014.

[13] H. Janetzko, D. Sacha, M. Stein, T. Schreck, D. A. Keim, and O. Deussen. Feature-driven visual analytics of soccer data. In *Proceedings of IEEE Conference on Visual Analytics Science and Technology*, pages 13–22, 2014.

[14] L. Jin and D. C. Banks. Visualizing a tennis match. In *Proceedings of IEEE Symposium on Information Visualization*, pages 108–114, 1996.

[15] M. Krstajic, E. Bertini, and D. A. Keim. CloudLines: Compact display of event episodes in multiple time-series. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2432–2439, 2011.

[16] M. Lage, J. P. Ono, D. Cervone, J. Chiang, C. Dietrich, and C. T. Silva. StatCast Dashboard: Exploration of spatiotemporal baseball data. *IEEE Computer Graphics and Applications*, 36(5):28–37, 2016.

[17] M. Lames. *Leistungsdiagnostik durch Computersimulation: ein Beitrag zur Theorie der Sportspiele am Beispiel Tennis*. PhD thesis, University of Mainz, Germany, 1991.

[18] P. A. Legg, D. H. Chung, M. L. Parry, M. W. Jones, R. Long, I. W. Griffiths, and M. Chen. MatchPad: Interactive glyph-based visualization for real-time sports performance analysis. *Computer Graphics Forum*, 31(3pt4):1255–1264, 2012.

[19] R. Leser and A. Baca. 30 practice oriented match analyses in table tennis as a coaching aid. *Science and Racket Sports IV*, page 214, 2008.

[20] Z. Liu, B. Kerr, M. Dontcheva, J. Grover, M. Hoffman, and A. Wilson. CoreFlow: Extracting and visualizing branching patterns from event sequences. *Computer Graphics Forum*, 36(3):527–538, 2017.

[21] T. C. Loh and O. Krasilshchikov. Competition performance variables differences in elite and u-21 international men singles table tennis players. *Journal of Physical Education and Sport*, 15(4):829, 2015.

[22] A. G. Losada, R. Therón, and A. Benito. BKViz: A basketball visual analysis tool. *IEEE Computer Graphics and Applications*, 36(6):58–68, 2016.

[23] I. Malagoli Lanzoni, R. Di Michele, and F. Merni. A notational analysis of shot characteristics in top-level table tennis players. *European journal of sport science*, 14(4):309–317, 2014.

[24] P. McLachlan, T. Munzner, E. Koutsofios, and S. C. North. LiveRAC: interactive visual exploration of system management time-series data. In *Proceedings of ACM Conference on Human Factors in Computing Systems*, pages 1483–1492, 2008.

[25] M. Monroe, R. Lan, H. Lee, C. Plaisant, and B. Shneiderman. Temporal event sequence simplification. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2227–2236, 2013.

[26] T. Munzner. *Visualization Analysis and Design*. A.K. Peters visualization series. A K Peters, 2014.

[27] M. L. Parry, P. A. Legg, D. H. S. Chung, I. W. Griffiths, and M. Chen. Hierarchical event selection for video storyboards with a case study on snooker video visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1747–1756, 2011.

[28] A. Perer and D. Gotz. Data-driven exploration of care plans for patients. In *Proceedings of ACM Conference on Human Factors in Computing Systems*, pages 439–444, 2013.

[29] C. Perin, J. Boy, and F. Vernier. Using gap charts to visualize the temporal evolution of ranks and scores. *IEEE Computer Graphics and Applications*, 36(5):38–49, 2016.

[30] C. Perin, R. Vuillemot, and J. D. Fekete. SoccerStories: A kick-off for visual soccer analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2506–2515, 2013.

[31] C. Perin, R. Vuillemot, and J.-D. Fekete. A Table!: Improving temporal navigation in soccer ranking tables. In *Proceedings of ACM Conference on Human Factors in Computing Systems*, pages 887–896, 2014.

[32] M. Pfeiffer, H. Zhang, and A. Hohmann. A markov chain model of elite table tennis competition. *International Journal of Sports Science & Coaching*, 5(2):205–222, 2010.

[33] H. Pileggi, C. D. Stolper, J. M. Boyle, and J. T. Stasko. SnapShot: Visualization to propel ice hockey analytics. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2819–2828, 2012.

[34] C. Plaisant, B. Milash, A. Rose, S. Widoff, and B. Shneiderman. LifeLines: Visualizing personal histories. In *Proceedings of ACM Conference on Human Factors in Computing Systems.*, pages 221–227, 1996.

[35] T. Polk, J. Yang, Y. Hu, and Y. Zhao. TenniVis: Visualization for tennis match analysis. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2339–2348, 2014.

[36] D. Sacha, F. Al-amoody, M. Stein, T. Schreck, D. A. Keim, G. L. Andrienko, and H. Janetzko. Dynamic visual abstraction of soccer movement. *Computer Graphics Forum*, 36(3):305–315, 2017.

[37] M. Stein, H. Janetzko, T. Breitkreutz, D. Seebacher, T. Schreck, M. Grossniklaus, I. D. Couzin, and D. A. Keim. Director's cut: Analysis and annotation of soccer matches. *IEEE Computer Graphics and Applications*, 36(5):50–60, 2016.

[38] M. Stein, H. Janetzko, A. Lamprecht, T. Breitkreutz, P. Zimmermann, B. Goldlücke, T. Schreck, G. Andrienko, M. Grossniklaus, and D. A. Keim. Bring it to the Pitch: Combining video and movement data to enhance team sport analysis. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):13–22, 2018.

[39] D. S. Tan, G. Smith, B. Lee, and G. G. Robertson. AdaptiviTree: Adaptive tree visualization for tournament-style brackets. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1113–1120, 2007.

[40] R. Vuillemot and C. Perin. Sports tournament predictions using direct manipulation. *IEEE Computer Graphics and Applications*, 36(5):62–71, 2016.

[41] T. D. Wang, C. Plaisant, B. Shneiderman, N. Spring, D. Roseman, G. Marchand, V. Mukherjee, and M. S. Smith. Temporal summaries: Supporting temporal categorical searching, aggregation and comparison. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1049–1056, 2009.

[42] S. Wenninger and M. Lames. Performance analysis in table tennis-stochastic simulation by numerical derivation. *International Journal of Computer Science in Sport*, 15(1):22–36, 2016.

[43] K. Wongsuphasawat, J. A. G. Gómez, C. Plaisant, T. D. Wang, M. Taieb-Maimon, and B. Shneiderman. LifeFlow: visualizing an overview of event sequences. In *Proceedings of ACM Conference on Human Factors in Computing Systems*, pages 1747–1756, 2011.

[44] K. Wongsuphasawat and D. Gotz. Exploring flow, factors, and outcomes of temporal event sequences with the outflow visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2659–2668, 2012.

[45] J. Wood. Visualizing personal progress in participatory sports cycling events. *IEEE Computer Graphics and Applications*, 35(4):73–81, 2015.

[46] Y. Wu, J. Lan, X. Shu, C. Ji, K. Zhao, J. Wang, and H. Zhang. iTTVis: Interactive visualization of table tennis data. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):709–718, 2018.

[47] H. Zhang, W. Liu, J.-j. Hu, and R.-z. Liu. Evaluation of elite table tennis players' technique effectiveness. *Journal of sports sciences*, 31(14):1526–1534, 2013.

[48] J. Zhao, S. M. Drucker, D. Fisher, and D. Brinkman. TimeSlice: interactive faceted browsing of timeline data. In *Proceedings of International Working Conference on Advanced Visual Interfaces*, pages 433–436, 2012.

[49] J. Zhao, Z. Liu, M. Dontcheva, A. Hertzmann, and A. Wilson. MatrixWave: Visual comparison of event sequence data. In *Proceedings of ACM Conference on Human Factors in Computing Systems*, pages 259–268, 2015.